NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**HACKING SOCIAL NETWORKS:
EXAMINING THE VIABILITY OF USING COMPUTER NETWORK ATTACK AGAINST SOCIAL NETWORKS**

by

Russell G. Schuhart II

March 2007

Thesis Advisor: David Tucker
Second Reader: Karl Pfeiffer

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2007	Master's Thesis

4. TITLE AND SUBTITLE: Hacking Social Networks: Examining the Viability of Using Computer Network Attack Against Social Networks	5. FUNDING NUMBERS
6. AUTHOR(S) Russell G. Schuhart II, LT, USN	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution is unlimited.	A

13. ABSTRACT (maximum 200 words)

Social Network Analysis (SNA) has been proposed as a tool to defeat transnational terrorist groups such as Al Qaeda. However, SNA is an descriptive tool that is a product of sociology and not an offensive tool used to attack a social network. SNA was not designed to destabilize covert networks that are difficult to observe and penetrate. This work presents a possible way to improve SNA's performance against a covert social network by employing the Computer Network Attack (CNA) model. The CNA model is used by computer network security to represent the traditional approach to hacking a computer network. Although not tested in this paper, it is argued that the CNA model should be able to improve the accuracy of SNA when applied to a covert social network by standardizing the destabilization process and allowing for frequent challenges of operating assumptions.

A history and overview of both computer networks and social networks is covered to allow for a comparison of the two networks. The networks have enough similarities to allow the application of the CNA model without major modification from its original form. Assumptions about the security of computer and social networks are examined to clarify how the CNA model can attack a social network. The model is examined for validity and the conclusion is that the CNA model can incorporate SNA into a more methodical approach to achieve better results that using SNA alone. The final portion of the paper details a possible implementation of the CNA model and how it can be used as part of an offensive effort to destabilize a covert social network.

14. SUBJECT TERMS Social Networks, Social Network Analysis, Computer Networks, Computer Network Attack, Hacking, Networks, Network Theory	15. NUMBER OF PAGES 73
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited.

HACKING SOCIAL NETWORKS: EXAMINING THE VIABILITY OF USING COMPUTER NETWORK ATTACK AGAINST SOCIAL NETWORKS

Russell G. Schuhart II
Lieutenant, United States Navy
B.S., United States Naval Academy, 2001

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS AND OPERATIONS

from the

NAVAL POSTGRADUATE SCHOOL
March 2007

Author: Russell G. Schuhart II

Approved by: David Tucker
Thesis Advisor

Lt Col. Karl Pfeiffer
Second Reader

Dan Boger
Chairman, Department of Information Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Social Network Analysis (SNA) has been proposed as a tool to defeat transnational terrorist groups such as Al Qaeda. However, SNA is a descriptive tool that is a product of sociology and not an offensive tool used to attack a social network. SNA was not designed to destabilize covert networks that are difficult to observe and penetrate. This work presents a possible way to improve SNA's performance against a covert social network by employing the Computer Network Attack (CNA) model. The CNA model is used by computer network security to represent the traditional approach to hacking a computer network. Although not tested in this paper, it is argued that the CNA model should be able to improve the accuracy of SNA when applied to a covert social network by standardizing the destabilization process and allowing for frequent challenges of operating assumptions.

A history and overview of both computer networks and social networks is covered to allow for a comparison of the two networks. The networks have enough similarities to allow the application of the CNA model without major modification from its original form. Assumptions about the security of computer and social networks are examined to clarify how the CNA model can attack a social network. The model is examined for validity and the conclusion is that the CNA model can incorporate SNA into a more methodical approach to achieve better results that using SNA alone. The final portion of the paper details a possible implementation of the CNA model and how it can be used as part of an offensive effort to destabilize a covert social network.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

The author would like to thank his family for their support and patience during the writing of this thesis. A loving wife and a new son have made this a happy time for all of us. They have made this effort possible.

The author wishes to thank Professor David Tucker for the space given to work without constant oversight, and for his ability to see statements presented as fact, that were based on very little that was factual. He has made this effort a success.

Finally, Lt. Col. Pfeiffer deserves acknowledgment for his constant willingness to assist students in any way he can. Specifically, the author thanks him for being the second reader to this thesis, but also for his help in becoming a part of the ISO major and his efforts from the author's first days at NPS.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. THE OBJECTIVE OF THIS PAPER

Social Network Analysis (SNA) is used by sociologists to measure and to understand the social networks that develop when people interact. The study of social networks has been part of sociology since the mid-1950s, and SNA has evolved and solidified slowly in the past fifty years. Recently, world events have thrust SNA into the spotlight. The Global War on Terror and the post 9-11 security environment have focused a great deal of attention on alternative uses for SNA. Global terrorist organizations such as Al Qaeda are organized into structures that are typical, fragmented social networks. The recent attention on SNA is centered on what it can bring to bear in defeating a terrorist group such as Al Qaeda, and what possibilities may exist in using SNA as an offensive measure to destabilize a network instead of an academic tool used solely to understand a network's relationships.

One of the hurdles to SNA making the change from academic device to operational asset is that SNA has yet to experience the growth and refinement customary for accepted use of an analytical tool. The standardization and predictability necessary for employment by forces battling a covert network does not exist. SNA is being used to fight covert networks, but its use is limited to improving situational awareness instead of integrating intelligence, analysis, and targeting to meet strategic objectives in defeating covert social networks.

If we wish to use SNA in the context of offensive operations, then it must be standardized into an offensive framework. A framework already in use is the CNA (Computer Network Attack) model created in computer network security to model computer hacking. The model is based on four stages that a typical computer "hacker" progresses through to penetrate a computer network: footprinting; scanning; enumeration; and exploitation. Each stage is offensive, builds on the previous stages, and is geared towards a clandestine penetration of a network. The analogy between using SNA to penetrate a social network and using the CNA model to exploit a computer network appears rather intuitive. However, to take the CNA model from analogy to a

viable method for improving SNA requires careful examination, and that is the purpose of this paper: to determine if the CNA model can improve SNA applied to a covert social network, and exactly how applicable the model is to destabilizing a covert network.

B. MEETING THE OBJECTIVE

Multiple assumptions must be challenged and verified before the CNA model can be evaluated. The first is that social networks and SNA will be examined to ensure that there is a clear understanding of what they are and how they operate. Just with any skill-set or tool, SNA comes with its own terminology and metrics. Only a basic introduction will be covered here, but enough to ensure that the reader understands how SNA is applied.

Once SNA is understood in a benign setting, the challenges of applying it to a covert social network will be detailed to allow for a later comparison with the CNA model. Current research in this area will be examined to highlight some of the issues of using SNA against a covert network. SNA is a descriptive tool designed to describe the relationships in a social network with various methods such as a graph or a matrix. However, SNA runs into many problems with data accuracy when facing a network that wishes to remain on the fringes of society. The analyst cannot survey the network members for their relationships, and historical records are of questionable accuracy. Therefore, SNA must make several assumptions when analyzing a network. These assumptions are potential sources of error in the final network analysis. Each assumption will be examined to determine how the CNA model can improve its accuracy.

After exploring social networks, computer networks and their operations are examined. Computer networks handle their communications using set protocols and procedures that are designed for reliability and predictability. There are only a few possible arrangements of computers in a network topology and they are normally configured by convention using a best-practices approach. Computer networks have a separate field of Network Security filled with professionals that keep networks safe and hackers out. The way that computer networks are secured is important to understanding how the CNA model is applied against computers.

Once social networks and computer networks are examined separately, they will be compared against one another. Social networks are composed of people that are inherently complex and unpredictable when compared to the computers in a computer network. Computer networks are established and organized in predictable ways following standardized guidelines. The two networks are similar in many ways, but are also different in an equal number of ways. The two different networks create a need to test the underlying assumption that applying the CNA model to a covert social network is possible. This assumption will be tested to determine if the CNA model needs to be modified, or if the model is applicable in its original form.

After the CNA model has demonstrated some validity in its application to covert social networks, the paper will question is ability to improve SNA. This is necessary because the model can be applicable, but still not improve SNA in any way. When using the CNA model, SNA will form a subset of a more complete process. In this way, it should improve the accuracy of SNA and ensure more reliable results from the analysis. Finally, the paper will conclude by merging the CNA model with SNA to demonstrate how the two can be intertwined to fight a covert network.

THIS PAGE INTENTIONALLY LEFT BLANK

II. SOCIAL NETWORKS, COVERT NETWORKS, AND SOCIAL NETWORK ANALYSIS

A. INTRODUCTION

Since 9/11 and the Global War on Terrorism, the analysis of social networks has become a subject of great interest. The attacks on the World Trade Center focused the public's attention on transnational terrorists groups like Al Qaeda. The need to understand and map terrorist networks requires unorthodox thinking. Fighting terrorist networks that are often organized into disjointed cells can present a challenge for even the most experienced military or law enforcement investigator. The study of social networks and Social Network Analysis (SNA) have emerged as a potential tool to aid in the pursuit and neutralization of terrorists. To understand the potential and limitations of social networks, it is necessary to examine what a social network is and is not, but also how SNA can be used to model and disrupt covert social networks.

B. SOCIAL NETWORKS

Social networks can be divided into two components: a social component and a network component. In general, when referring to social networks, the social component is the existence of some relationship between two or more people, places, or events. It could be a friendship, club membership, school, or any other connection that unites the members in some social relationship. The second component, the network, is the grouping of the people, places, or events into an interconnected organization that has no formal or hierarchal construction. The network is characterized by a flat organization of interconnected members in which most member have equal status or rank. Formally, a social network is defined as: "a finite set or sets of actors and the relation or relations defined on them" (Wasserman and Faust, 20).

Even when formally stated, the definition in practice varies greatly from researcher to researcher, and the term is often used loosely. However, interconnection, a lack of formal structure, and relative equality among members are the common characteristics. It is also important to note that some authors draw a distinction between a network and the "flat organization" with only a few leaders who exercise control over a

number of subordinates of the same status or rank (Fukuyama and Shulsky, 1997). This distinction will not be drawn here. Networks and flat organizations will be treated as one and the same.

C. A BRIEF OVERVIEW OF SOCIAL NETWORK ANALYSIS (SNA)

1. The Motivation behind SNA

SNA is the study of social networks. Studying social networks requires a separate approach from traditional sociological methods. Traditional sociology focuses on the actor and how the actor's attributes (race, gender, ethnicity, etc.) affect the data in question (Degenne and Forse, 1). However, when examining social networks, the attributes of the actor are secondary to the connections between the members of the network. Sociologists were forced to adapt their thinking to study the *interconnection* of people instead of studying *people* when they are interconnected. The difference may appear subtle, but it is not. The focus is on the relationship instead of the people in the relationship. This change of focus lead to the merging of network theory from other fields, such as mathematics, with accepted practices from sociology. The result of this merger was Social Network Analysis.

There are two major characteristics that are unique to SNA. The first is that SNA focuses on the relationships amongst the network members. The second characteristic is that the actors are assumed to be interdependent instead of independent (Wasserman and Faust, 4). As they are interdependent, metrics are compared against the network or portions of the network. Individual members of the network may be compared to one another, but the metrics used to compare the nodes will derive their value from comparison against the network as a whole. For example, one node may be more connected than another node, but the metric of connectivity takes on meaning because it measures how a single node relates to the entire network. Individual properties of nodes (race, gender, etc.) are not the primary concern.

2. Terminology

SNA comes with a long vernacular that is varied and colorful. There are often multiple terms that refer to the same concept, and these differing terms can lead to confusion. Establishing a few terms is necessary to standardize a common vocabulary. *Actors* or *Nodes* refer to the members of a social network. Some social networks measure

the relationships between people (e.g., friendship) and others measure the relationship between people and some common event or location (e.g., organization membership or party attendance). *Tie* is used interchangeably with relationship to express the shared bond among actors or nodes.

The term *mode* is frequently encountered in SNA. The concept of a mode is easy to understand, but can be difficult to explain. Wasserman and Faust define mode as: "a distinct set of entities on which the structural variables are measured." (29) Mode refers to the "set" of actors. If all the actors come from the same set (group, religion, party, school, etc.), then it is a *one mode network*. If the set of actors come from two different groups (two different political parties, groups, teams, etc.) then it is a *two mode network*. Note that an actor/node can be an event or a location. If a group of people are meeting at a location, it would be a two mode network. The first mode is the group of people and second is the location or the meeting itself. Any number of modes beyond two is generally referred to as a *multimode.*

3. Matrices and Graphs

SNA's greatest potential lies with its ability to display relational data. This is most commonly done as a sociomatrix or as a sociogram. Both sociograms and sociomatrices have been in use since the 1930s and are a product of the science of sociometry which studies affective relations among actors such as like/dislike or love/hate (Wasserman and Faust, 77). The more advanced concepts of SNA derive their methods and importance from the use of either a sociomatrix or a sociogram.

A sociomatrix is a matrix that shows the relationships among different actors. Each axis of the matrix contains the actors/nodes in the network. The cells of the matrix show the relationship between the actors. Common coding schemes are binary and valued. Binary coding is used to show that a relationship exists (1) or a relationship does not exist (0). Valued coding is used to show different degrees of relationships such as acquaintance (1), friend (2), or best friend (3). Table 1 below provides an example of a two mode binary sociomatrix and reflects who attended a party thrown by Jim. Table 1 also shows a one mode, valued sociomatrix that reflects the friendship coding scheme mentioned above.

	Jim's Party
Mike	1
Bob	0
Jen	1
Tim	1
Deb	0

Two mode, binary

	Mike	Bob	Jen	Tim	Deb
Mike	-	2	2	3	2
Bob	2	-	1	2	1
Jen	2	1	-	2	2
Tim	3	2	2	-	2
Deb	2	1	2	2	-

One mode, valued

Table 1. Examples of SNA Coding Schemes

Sociograms improve visualization of network composition and the arrangement of actors. They are most often graphs composed of actors/nodes represented as points that are connected to one another by lines that represent relationships/ties. Most frequently, sociograms use sociomatrices as their input data. The data from the matrix become the coordinates in a Cartesian plane, or more simply, the matrix data becomes the X and Y coordinates for the sociogram when plotted. Below is an example of sociograms created in the NETDRAW software package:

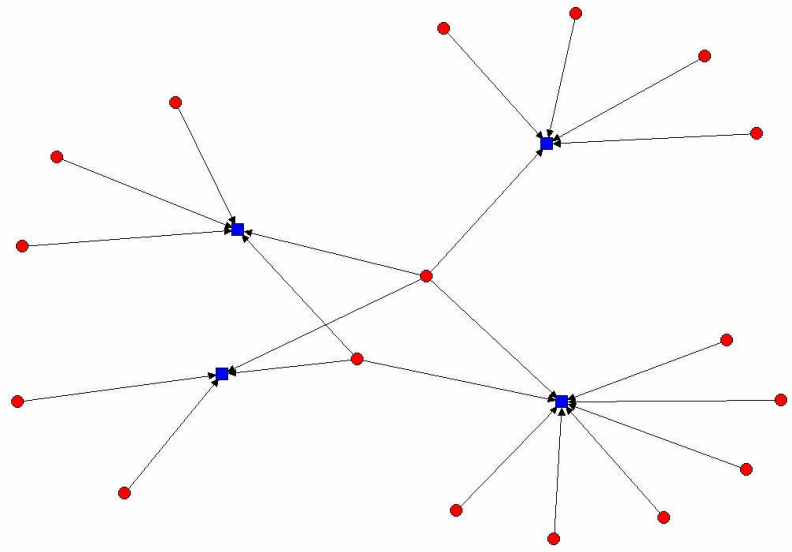

Figure 1. Two Mode, Directed Sociogram

The sociogram above displays directed relationships. Sociograms can be drawn as directional or as non-directional. The difference is that directional sociograms use arrows to indicate the direction of relationships between nodes. This is useful for indicating relationships such as like and dislike, or more relevant to covert networks,

which actors are supplying other actors or which actors seek information from other actors. Non-directional ties are used in situations where there is no need to denote choice or affective relationships. For example, a sociogram displaying the membership of a drug cartel would use simple lines instead of arrows to connect nodes, as membership is an absolute metric without a need to indicate direction.

Often analysts will break a sociogram into smaller pieces to prove a point or to highlight a unique characteristic of a relationship. There are three general levels of data that can be analyzed and are often referred to as the "modeling unit" (Wasserman and Faust, 44). The first modeling unit is the *dyad* which is the grouping of only two nodes. The second level or modeling unit is the *triad* which is the collection of three nodes. The final modeling unit is the *subgroup* or *subnetwork*. A subgroup is a collection of nodes that can range from two nodes to one node less than the original graph. Modeling units are necessary for discussing network patterns and metrics and are an important part of applying SNA.

4. Data Collection and Measurement

Social network data is traditionally collected by interacting with network actors or observing their relationships. Surveys, questionnaires, observation, and experiments are the methods normally used. However, when there is no way to interact with the actors or to gain access for observation, as in a terrorist organization, analysts must use archival data or public records to assemble the social network. A classic example of using public records to construct a social network was the reconstruction of the September 11th hijackers by Valdis Krebs in his "Mapping Networks of Terrorist Cells." Krebs used information from major newspapers, released data about the relationships amongst the hijackers from law enforcement agencies, and Internet search engines to determine how the network interacted. Krebs's study is proof that social network data collection can require creativity when the actors are not accessible. Data collection and manipulation requires careful definitions and qualifications to prevent a study from sliding from logical conclusion into conjecture or presumption.

Social network metrics are always measures that compare a node or a tie to the other nodes and ties in the network. To study relationships the collected data cannot be viewed in isolation, and it only takes on meaning when it is viewed in relation to the

larger dyad, triad, subnetwork, or the network as a whole. For example, a SNA metric is *centrality.* Centrality comes in different forms such as *degree centrality* which is the number of connections to other actors or *betweenness centrality* which is a measure of how many ties pass through an actor. In general, those nodes that form hubs of the network are considered central. Researchers intuitively understand the concept and have also found it "difficult to quantify," but agree that power concentrates with central actors (Degenne and Forse, 132). Centrality relates an actor to the other nodes in the network. The fact that SNA metrics are only of value when compared to a larger network is important to understanding what a social network analyst is attempting to prove and also to understanding what assumptions, errors, and mistakes may have occurred in data collection or processing.

In summary, Wasserman and Faust in their classic book, *Social Network Analysis*, provide an excellent overview of SNA:

> Given a collection of actors, social network analysis can be used to study the structural variables measured on actors in the set. The relational structure of a group or larger social systems consists of the pattern of relationships among the collection of actors. The concept of a network emphasizes the fact that each individual has ties to other individuals, each of whom in turn is tied to a few, some, or many others, and so on. The phrase "social network" refers to the set of actors and the ties among them. The network analyst would seek to model these relationships to depict the structure of a group. One could then study the impact of this structure on individuals within the group. (9)

5. Patterns and SNA

Of great importance to SNA is the central role that patterns play in understanding a social network. What is apparent in the summary of SNA provided by Wasserman and Faust is that patterns are the key component that turns raw data into useful information. They refer to the emergence of patterns as structure, and point out that the analyst must model this structure to understand the network and gather useful information such as composition, organization, or roles.

The single key that makes SNA plausible and useful is that interacting people form patterns. SNA reverses the interaction of people to determine why a network exists in its present form. People meet one another and then based on some rationale (i.e.

shared interests, sense of humor, etc.) they form patterns of interaction (i.e. seeking advice from the same people, eating lunch with the same group, etc.). SNA reverses the process by modeling the interactions between people to determine a pattern, and then determines why the actors created these patterns.

Network analysts assume that there is always a pattern; otherwise the network cannot be studied. A pattern is something interesting that emerges amongst the connections of nodes with links in either a matrix or a graph. Within the connections, some nodes will pique the interest of the analyst because they are more connected, more central, or for some other reason. The descriptive capability of SNA helps an analyst find a pattern. Without a pattern, there would be no network to study. The challenge for the analyst is to keep working with the data and to change the model, also known as *permutation*, until a pattern emerges. Therefore, pattern recognition and pattern interpretation are the two pillars of successful SNA. Once the pattern is found (pattern recognition) the analyst can then determine how to change the network (pattern interpretation) to obtain the desired effects.

D. SNA AND DISRUPTING COVERT NETWORKS

SNA is traditionally used to understand sociological questions about friendships, work relationships, and other common, everyday applications where actors organize into informal relationships. However, recently the idea of using SNA to track and disrupt terrorist networks has gained traction. In fact, the newly released joint U.S. Army and U.S. Marine Corps counterinsurgency manual, FM 3-24, includes an appendix that addresses the basics of SNA (E-10).

SNA has the potential to be a powerful tool for tracking and identifying obscure patterns in a terrorist or criminal network, but its applications are limited. SNA can only provide the analyst or warfighter another way to describe the enemy's organization. SNA cannot topple Al Qaeda and it cannot bring a terrorist network to its knees. Current efforts to use SNA as a destabilizing force lack depth and structure. In general, SNA is too immature to be actively used in the way some researches are proposing, and is forced to remain as a useful tool for improving the commander's situational awareness and understanding of an enemy network.

For the sake of clarity, in the discussion that follows "covert networks" or "covert social networks" are defined as networks that engage in illegal activity and attempt to operate clandestinely. Other authors have used the same criteria to define "dark" and "light" networks (Rabb and Milward, 415).

E. THE CHALLENGE OF COVERT NETWORKS

SNA is normally used in a benign manner to track social networks. Its methods and techniques encounter major challenges that require some rather drastic changes when applied to covert social networks. The fact that covert networks operate with utmost secrecy and are organized to resist detection or penetrations are major obstacles for the analyst applying SNA. Although these challenges are not only faced by SNA, but also by any analytical study of a covert network, SNA's ability to describe the network in question is severely hampered when studying a covert network. Data collection, data accuracy, and pattern detection all differ from traditional SNA.

Terrorist networks' survival is directly proportional to secrecy. Skilled terrorist groups organize into cells to prevent discovery and limit damage if cells are captured by authorities. Traditional data collection methods like surveys and interviews are not applicable to terrorist networks for obvious reasons. SNA must fight against this secrecy to determine which actors are part of the network and what roles they fill. Analysts are forced to rely on intelligence reports, interrogations, and logical deduction as well as educated guessing to gather network data. Invariably, this leads to errors. Some actors may escape detection and other nodes will wrongly be assigned roles and positions to compensate. The model network structure may or may not accurately reflect the actual terrorist network. There is no way for the analyst to know how accurate the model is. It cannot be overstated that every sociogram and sociomatrix is nothing more than an educated guess with varying degrees of probabilistic certainty.

SNA also faces a constantly changing data set. Every day, new information is collected while counter-terrorist and counterinsurgency efforts are ongoing. The analyst must be able to quickly modify the network model to reflect new data. Not only is data constantly changing, the covert network is also continuously adapting. Nodes are added

and removed, change position, and change importance as the network operates. This produces yet more error in the model network. Covert social networks can change so quickly, that by the time the network data is gathered, the actual network may have changed drastically. Overt networks also change rapidly, but the openness of the network should allow the analyst the ability to adjust his data and models much more quickly that changes in a covert social network. Frequently studies of covert networks are well after the fact and are used for illustrative purposes instead of attempting to actually destabilize standing networks.

Krebs found in his study of the September 11th hijacking that he faced the "inevitable" problems of network adaptability and data completeness (44). There is no way to overcome these challenges. As their effects cannot be removed, they must be mitigated. It falls to the analyst to ensure that the data set is as complete as possible, but also to determine a good faith estimate of how accurate the data is. SNA can be applied to covert networks, but it must be done carefully and with a great deal of respect for its limitations.

F. CURRENT RESEARCH ON DESTABILIZATION

Researchers that are applying SNA and network analysis to covert networks are few in number. Krebs and the 9/11 study have already been mentioned. Jonathan Farley examined the Al Qaeda network from a mathematical point of view (Farley, 2003). However, the most prolific writer and researcher has been Professor Kathleen Carley at Carnegie Melon University. She has written numerous papers on mapping, destabilizing, and understanding covert networks. She has been writing about networks for just shy of two decades, and the fact that she frequently sites her own work is indicative of the lack of research in applying SNA to covert networks. Unfortunately, Carley has focused more on developing software and methods to compare destabilization strategies than actually proposing how SNA can be used to fight a covert network.

Carley's papers have repeatedly presented the same strategy for destabilization (Carley, 2001, 2003, 2006). Her strategy is to add or remove nodes, but the focus is on removal. This is also the strategy proposed by Farley. A node is targeted using SNA and her evolution of SNA termed Dynamic Network Analysis (DNA). SNA and DNA will be treated as the same tool. Although she proposes that DNA is different from SNA,

from her papers it appears that DNA is nothing more than software and computer processing applied to SNA. However, according to Carley:

> Dynamic network analysis (DNA) combines social network, statistical link analysis, machine learning, artificial intelligence and computer simulation techniques to create systems for estimating the size, shape (Dombroski et al., 2003), vulnerabilities, key actors and dynamics of relational data (Carley, 1999) connecting various entities including people, resources, actions and locations (Carley 2006, 54).

Carley (2006) brings together all of her previous work into one document and specifically addresses covert networks. Nodes are targeted for removal based upon their centrality, possession of a unique skill (such as a bombmaking), or their high "cognitive demand" which is requirement for an "emergent leader" (Carley 2006, 56). Carley points to her earlier work to explain that emergent leaders are those nodes that display high "cognitive load" and are worthy of targeting.

> Overall cognitive load, not simply structural power, is key to tracking who is likely to be the emergent leader. Based on these considerations, we define the emergent leader as the individual with the highest cognitive load (the most people to talk to, the most information to process, the most tasks to do, the hardest tasks to do, the most people to negotiate with to get the job done, etc.)…Consequently, emergent leaders, by virtue of their centrality across the entire meta-network are good candidate agents to remove if the goal is to destabilize the network (Carley 2001, 84).

Once these nodes are removed, Carley argues, the network will suffer from "cascading errors" that will reduce the network's ability to function. Unfortunately, Carley (2006) does not use a real world example to prove her case, but instead uses a greatly simplified, virtual experiment conducted with software to isolate nodes in a network in succession. She compared random networks and cellular networks and also compared the results of isolating emergent leader nodes, central nodes, and random nodes. Her conclusions from the experiment are that cellular networks are more susceptible to nodal isolation or what she terms an "attrition strategy" than random networks and "The results also indicate that for cellular networks they can be destabilized best, in the short run, by isolating those individual's [*sic*] who are emergent leaders" (Carley 2006, 66).

Carley (2006) presents some interesting thoughts, but also gives cursory mention to the challenges that SNA faces when used against covert networks. To employ a strategy of nodal removal, the structure of the network must be known. Determining who to remove requires quite a bit of intelligence. The reality of fighting covert networks is that this information is often not available due to the covert network's secrecy and rapid adaptation. In a protracted destabilization effort, the required information may be gathered and applied, but nodal removal is a simplistic approach to destabilization. The United States has been targeting Osama bin Laden for over five years and has been unable to remove this node. It is debatable how central he his to Al Qaeda's operations, but the point remains that nodal removal is not as easy in practice.

The repetition of nodal removal is not the only example of researchers making simplifying assumptions that neglect the reality of fighting covert networks. SNA functions on pattern recognition and pattern interpretation. The detection of patterns in the past is easy as the data is known, and often prior patterns can provide insights to the future. Krebs study of the 9/11 hijackers was successful because it occurred after the fact, and is useful for understanding how the network organized and operated. His study is not directly applicable to the future and it certainly cannot be assumed that the next attack on America will operate in the same way as the 9/11 network. SNA must use known terrorist networks to demonstrate its abilities against covert networks. However, this trend also induces an inherent focus on the past. The regularity that is required for pattern detection is difficult to obtain in an operating covert network. Even with the possible errors that arise when applying SNA to covert networks and the challenges of data integrity outlined previously, SNA is not a predictive tool and attempting to use it as such discounts too many variables to give results any degree of validity.

G. WHAT SNA IS MISSING

SNA may become an important operational tool as the Global War on Terrorism remains a major component of the National Security Strategy. The fact that recently published counterinsurgency manual, FM 3-24, includes SNA is proof that there is belief in the merit of its application. However, the current state of research in SNA and its use against covert networks presents serious questions about its applicability and utility. The

problems with data collection have already been covered in detail. It is understandable why SNA is met with skepticism, and SNA must overcome these limitations and address what is missing if its benefits are to be realized.

Destabilization strategies should include more than nodal removal. The absence of other presented strategies is disconcerting, especially when law enforcement and intelligence agencies have been practicing SNA, although in less codified form, since its inception. Nodes can be neutralized by killing or kidnapping them as Carley (2001, 2003, and 2006) and Farley (2003) suggest, but they can also be manipulated. Carley discusses adding nodes to a network, but dismisses it as being impractical and "a perilous and slow strategy" (Carley, Destabilization 58). Adding nodes is difficult, but it should not be dismissed because it is not a quick solution to network destabilization. Additionally, current research neglects that the manipulation of covert networks can often produce more desirable and predictable results than attempting to fragment the network or causing "cascading errors" using nodal removal.

Leaving actors in place presents opportunities to turn them into agents for friendly forces. Intelligence agencies have perfected this to an art, and their art form is ideal for manipulating the relationships between actors. Although recruiting and running agents is a complex endeavor, even basic tradecraft such as appealing to any of the Seven Deadly Sins (greed, lust, pride, etc.) should produce benefits. Unreachable nodes can be accessed by using more exposed nodes and turning them into agents. As well, the primary goal of any analyst should be to obtain as much information as possible in order to create the best possible match between the model network and the actual network. Leaving nodes in place allows information about the network to expand and the probability of error decreases as actors are under surveillance.

The information that is passed between actors can be manipulated to draw suspicion or to create confusion. False information spread with a careful touch can create rumors and distrust to manipulate the network. For example, planting information that an actor is meeting with friendly forces can deteriorate the trust that exists in the network ties. An appealing aspect of this strategy is that the other members of the network may move from suspicion into action and possibly neutralize a node that cannot be reached by

friendly forces. These are only two alternatives to nodal removal and future research should focus on other options.

The options for network manipulation are only limited by the creativity of the analyst and the operational forces employing SNA. SNA shows the analyst the patterns in the network and nothing more. SNA will not present the best possible method to destabilize the network, it is a descriptive tool and it can only show the network's structure and the patterns that can be destabilized or manipulated. Often proponents of SNA will forget the fact that SNA is primarily a tool for analysis and when applied for forecasting or prediction it quickly builds up errors that leave it little more than guesswork.

1. A Systematic Process

The lack of alternative destabilization strategies points to the lack of structure or formal process in applying SNA to covert networks. While research is lacking on the whole, no formal, systematic process exists for attacking and defeating covert networks. Using a systematic approach allows all users to examine data collection and proposed strategies from start to finish. A formal process would help increase the confidence level of model networks because analysts would all use the same procedure to map a network. Analysts not involved with the original study could examine the network analysis and quickly determine where assumptions were made and if the assumptions were reasonable. The potential for collaboration among analysts will certainly produce more accurate and reliable network models.

Instead of creating a process for attacking a covert network, it is more efficient to apply a process that already exists. This process exists in computer networks and Computer Network Attack (CNA). A systematic process exists for identifying the network structure, properties of a node, and vulnerabilities for exploitation. Exploiting the covert network still requires the creativity of the analyst, but adopting the process can help standardize SNA. The goal is to increase the confidence level of models and eventually lead to more successful destabilization efforts against covert networks.

THIS PAGE INTENTIONALLY LEFT BLANK

III. COMPUTER NETWORKS, NETWORK SECURITY, AND COMPUTER NETWORK ATTACK

A. INTRODUCTION

Most people have an intuitive understanding for of social networks. Friends, family, and coworkers are only a few of the thousands of possible networks that exist for every member of society. These networks serve as personal experience and are the basis for intuitive understanding. What is not necessarily intuitive to the average person is the way that computer networks are structured and how they operate. As will emerge in this chapter, there are many similarities between computer networks and social networks. The utility of these similarities will be explored in the next chapter, but this chapter aims to explore computer networks and the Computer Network Attack (CNA) model. The purpose is to enable a comparison between social networks and computer networks, and to highlight the benefits that CNA can bring to Social Network Analysis (SNA).

Before understanding the CNA process and framework, it is necessary to understand the way computers communicate, are arranged into networks, and how computer networks are secured. There are formal rules and guidelines that cover the interconnection and communication of computers. These rules and standards were designed for simple and reliable operation in order to meet the needs of legitimate users. However, these standards were not designed to stop, or even prevent, the abuse of a network by malicious users. "Hackers" and "crackers" are often used to describe those that find holes in the standards of communication and abuse their weaknesses for personal gain. The field of Network Security has responded to hackers' efforts to undermine the safe and secure operation of computer networks by evolving into an established, professional field. This evolution has prompted professionals to develop a formal model of hackers' approach to attacking a computer network. This model is the CNA framework and it can help improve the accuracy of, and confidence in, social network analysis.

B. COMPUTER NETWORKS

1. Characteristics, Operation, and Standardization

The first requirements for successful communication between two or more computers are standardization and openness. The means by which a computer transmits and receives data must be known by all other computers on the network. The process must also be done in a standard way. Whatever process is used by one machine must be used by other machines, or at a minimum, it must be known by other machines so that the process can be reversed to regain the data in question. For example, if a computer sends an email message by breaking it into ten equal sections, numbering them sequentially, and sending each piece individually, the receiving computer must know to look for ten sections, to place them in the proper order, and then display the information to the user. If the process was done differently by each user or was done in a secret way, other users would be faced with an influx of data devoid of meaning or structure. It is the same thing as two people attempting to communicate by speaking languages that they personally created. It simply will not work. There have to be known rules for how information is structured and passed between users.

Actual computer networks communicate with standardized processes that are formalized by different regulatory bodies. Although computer networks can be managed in any way by any user, to prevent confusion and possible damage, these groups act as the central reference for operating procedures. This prevents new technologies from conflicting with existing ones. The Internet Engineering Task Force (IETF), the Internet Assigned Numbers Authority (IANA), and the Institute of Electrical and Electronics Engineers, Inc. (IEEE) are a few of the major players in the standardization of computer networks.

These organizations decide the best way to deal with various questions such as: how data will be split apart and reassembled, which addresses can be used by different computers on the network and the proper format for those addresses. They deal with very basic questions about voltages and types of cables or connectors used to link computers. This standardization is a major hurdle for different technologies reaching mainstream usage. A government or private corporation may develop its own special software, but to

gain widespread use it may be modified to meet the standardization guidelines. In its original, proprietary format, software often prevents its incorporation into other applications. Once standardized, developers and customers know what to expect in terms of performance, interfaces, and other factors as all future instances of the software or technology will meet the assigned standards. Standardization allows technologies to take off and achieve the primary goal of generating income.

At the most general levels, computer networking follows a model created by the International Standards Organization (ISO) (Figure 2). This model is the Open Standards Interconnection Model, but is commonly referred to as the OSI Reference Model. The model itself has seven "layers" that represent a necessary step in the process of taking human communication (written, spoken, or visual) and reducing it to a machine usable format.

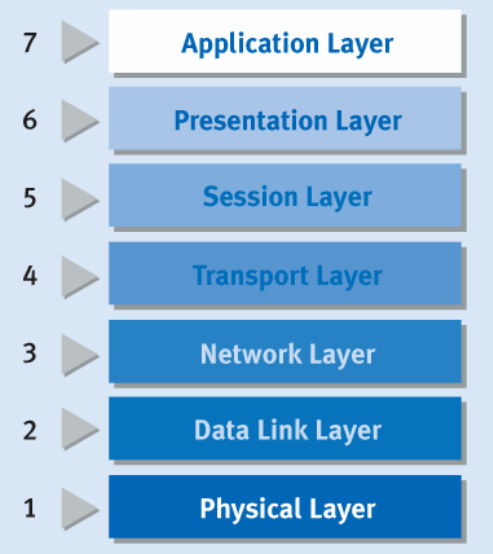

Figure 2. ISO Reference Model
(From: www.dell.com, Jan 16, 2007)

The important point about this model is that it is the definitive cornerstone of computer network communication. The OSI model is also extremely general. Real network communications can combine layers or modify the layers of the OSI model. As technologies emerge and mature, their respective developers will seek to innovate and will develop their own models that are similar to the OSI model. However, every

instance of computer intercommunication must incorporate all of the seven layers in some way. When different agencies and organizations create their own answer to fulfilling the requirements of the OSI model their solution is generally known as a *protocol suite*. A protocol is "a set of rules and formats that enables the standardized exchange of information between different systems" (Harris, 964). A suite is the collection of different protocols that are used in conjunction to meet the requirements set forth by the OSI model.

Without a doubt, the most common protocol suite is the Transmission Control Protocol/Internet Protocol (TCP/IP) suite (Goleniewski, 247). This suite is the foundation of the world wide Internet. The TCP/IP suite or "stack" was developed by the Department of Defense with the precursor of the Internet known as ARPANET. TCP is the protocol that handles the setup of connections, error control and correction, and is the "protocol for sequenced and reliable data transfer" (Goleniewski, 248). TCP functions as the intelligence in the suite, but IP is the workhorse of TCP/IP. IP handles basic functions such as breaking the data into smaller and more manageable sized units called "packets" or "datagrams," ensures packets are a set size with correct formatting and sequencing, and generally "handles packet forwarding and transporting of datagrams across a network" (Goleniewski, 248). TCP and IP are the means by which computer networks share information. There are other suites, but the TCP/IP suite is the dominant suite by far.

TCP/IP does not follow the ISO model exactly. There is a separate TCP/IP model, but it incorporates all of the layers of the ISO model. Communication for the ISO model or the TCP/IP model depends on the process of encapsulation. At a basic level, encapsulation is the process where data is broken into packets and information is added to the data as it proceeds down through the layers of the ISO or TCP/IP model. Each layer takes the previous layers' information and adds its necessary data to the outside of the packets. The layers are responsible for telling the network necessary information like who should receive the packet, what kind of information the packet contains, and how the receiving computer should reassemble the packets. The receiving computer reverses the

encapsulation to remove the required information from the outside of the packets, working its way to the original data, as the packet proceeds up through the layers of the model.

Defined as standard by the TCP protocol, each computer has 65,536 or 2^{16} possible "ports" that act as avenues of communication. Each port is akin to a door that the computer can communicate through to reach other computers. There is nothing that requires each port to be assigned to a protocol, but the IANA and accepted convention have firmly set many of the lower number ports. For example, Internet traffic such as web pages normally communicates on port 80. This allows any user desiring to determine if a machine has a web page up and running to examine port 80. If port 80 is "open," it implies that there is a webpage open for access.

Computer networks communicate based on the ISO model primarily using TCP/IP and defined ports to transfer information. However, one final characteristic is important to note. Computer networks can operate in two different modes: synchronous and asynchronous. In synchronous communication, the communicating computers are coordinating their transfer in some manner. Frequently, this synchronization is done via timing and special packets. Without coordination between the networked computers, they could not carry on a simultaneous conversation with one another because they could attempt to communicate at the same time and create conflicts on the network. Asynchronous networks are used for functions where only one computer will communicate at a time, such as broadcast traffic. Synchronous networks are more useful, but they are also more complex and present more opportunities for abuse.

2. Internetworking and the Internet

Computer networks can vary in size from two computers to an infinite number of machines. The size of the network will dictate the types of hardware and software used to communicate. A network that is designed for a small business of five users is vastly different from a business network used by hundreds of employees in different parts of the world. Connecting different computer networks into larger networks creates what is known as an internetwork or an internet. The largest example of the interconnection of networks is the global Internet that has become an integral part of daily life. Whether a small network or the Internet, when discussing computer networks the scale of the

network must be stated and understood as characteristics and conclusions drawn at one level may not be applicable to networks of different sizes.

There are two categories that are used to describe computer networks. There are more technical methods of division, but the categories are based on the geographic areas that they encompass. The first is the Local Area Network or LAN. LANs normally span a small area such as a campus or a single building. A university network is a classic example of a LAN. The second category is the Wide Area Network or WAN. WANs normally span multiple installations spread across greater distances such as counties, states, or countries. Using the university network again as a baseline, a WAN would be the interconnection of multiple university networks. WANs can be unlimited in size. The global Internet is an example of a WAN spanning the entire world.

WANs and the Internet have become areas of interesting research in recent years. This is an example of how LANs and WANs differ and why scale cannot be ignored when discussing computer networks. WANs are the subject of research instead of LANs because as networks grow in size, they also lose the unified oversight that can limit growth and expansion. Instead of a single organization having total control over the network, the involved networks must work in cooperation and with more permissive guidelines. These characteristics have allowed the explosive growth of the Internet and are the reason that researchers such as Albert-Laszlo Barbarasi have used the Internet to model virus growth, biology, social networks, economics, and other areas that incorporate network theory. Barbarasi found that the Internet has gown to the point that is almost is its own separate entity, “While entirely of human design, the Internet now lives life of its own. It has all the characteristics of a complex evolving system, making it more similar to a cell that to a computer chip” (Barabasi, Linked, 149). Barbarasi understands that the Internet is a collection of inter-networked computers and computers only do what they are told, but his point is that the Internet’s rapid growth and expansion parallel other natural processes such as cellular growth. The different scales of networks are drawing interest for their potential modeling applications. With this research comes the realization that LANs and WANs may be similar in construction and behavior, but their possible differences require that the scale of the network in question be stated before it is used for comparison, study, or discussion.

3. Predictability and the Lack of Options in Network Design and in Operating Systems

There are only a few topologies available for a network administrator when creating a LAN or WAN. The possible arrangements of machines are limited by the equipment in use and the purpose of the network. Along with limited topology is the limited number of computer operating systems (OSs). The easily enumerable options in topology and in operating systems leads to predictability in network constructions and operation.

There are only four major topologies that exist for computer networks. LANs are the best way to understand the topologies, but WANs follow the same principles, just on a larger scale. The four topologies are: bus, ring, star, and tree. Bus networks have all the machines connected to the same wire and are the most basic networks. Every message sent on the wire is seen by all other machines. A ring topology is similar to a bus except that the wire that connects the machines loops back on itself to form a ring. Ring is useful for applications where two loops are laid in close proximity to one another for redundancy. Star networks have computers all connected to a central piece of hardware such as router or a switch. Routers and switches allow the network to be segmented and provide more options to the designer. Tree networks have shared connections to a main data source like a bus network, but subnetworks branch off of the main source like the branches of a tree, hence the term "tree network." Mesh networks are a final topology that is not one of the major topologies, but may well be in the near future. Mesh networks have come into the spotlight with the growth of wireless networking. In a mesh network, every computer is connected to every other computer. All computer networks follow one these topologies or a combination of them. Figure 3 depicts the different topologies.

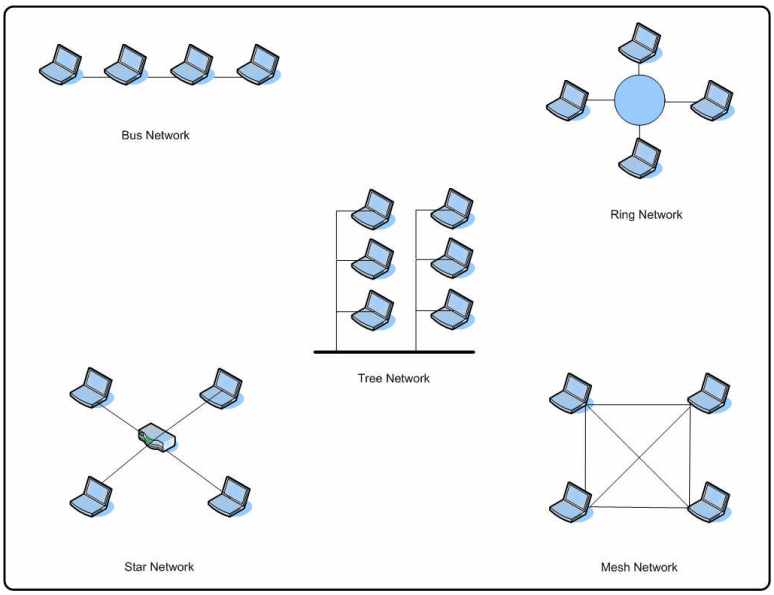

Figure 3. Basic Network Topologies

The final component of the predictability in computer networks is that there are a limited number of operating systems for the individual machines on the network. The most prevalent and common OS is Microsoft Windows. Significant, but not as prevalent are Apple's Mac operating system, as well as UNIX and its derivatives such as Linux, BSD, and Sun's Solaris. Other specialized operating systems Ss exist but are only used for special purposes and are insignificant in number when compared to the systems already mentioned. Each OS has its own "fingerprints" and patterns in the way it functions. How data is stored, written, transmitted, etc. can frequently be used to determine the type of OS in use. This is useful for network operations because it allows administrators to remotely work with different OSs and different machines, even if they have never actually seen the computer in question. However, the enumerable nature of OSs also limits the possibilities facing hackers and can make attacking a computer an easier proposition.

C. NETWORK SECURITY

The challenge of network security is that the extreme openness and predictability necessary for computer networks to operate efficiently works directly against the desire to keep network operations secure. The more that is known about network operations, the greater the likelihood that an attack will be successful. The protocols and systems developed as networks matured are not designed to stop hackers. There are many places

where the network assumes that commands and instructions are given by trusted personnel and does not check the authenticity of the user issuing commands. The openness that allows networks to function represents their major vulnerability.

1. Risk Management

If the way in which LANs and WANs communicate must be known to ensure reliability and functionality, then the goal of network security cannot be absolute security. If it were, then no traffic would be allowed to pass on the network out of fear of a possible attack. There is a spectrum of protection that has absolute security on one end (no traffic is allowed to pass), or on the other end, a total lack of security (allowing users and computers to do anything they choose). The goal is risk management and to minimize the risk as much as possible while still allowing the network to function while not exceeding other constraints such as budget and manpower. There is a difference between risk management and risk elimination. In her guide to gaining a network security certification, Shon Harris explains why network security's core is risk management:

> Information Risk management (IRM) is the *process* of identifying and assessing risk, reducing it to an acceptable level, and implementing the right mechanism to maintain that level. There is no such thing as a 100 percent secure environment. Every environment has vulnerabilities and threats to a certain degree. The skill is in identifying these threats, assessing the probability of them actually occurring and the damage they could cause, and then taking the right steps to reduce the overall level of risk in the environment to what the organization defines as acceptable (Harris, 65).

2. The Strategy of Network Security and the CIAA Framework

Network security and information assurance in general, are centered on a core framework that ensures the protection of the confidentiality, integrity, availability, and authenticity (CIAA) of information. These four pillars are often referred to as the CIAA model or framework, and the same convention will be used here. Every effort in network security will attempt to protect at least one of the components of the CIAA framework.

Confidentiality is the "assurance that information is not disclosed to unauthorized entities or processes" (DOD 8500.1, 18). Confidentiality in a network is normally provided by encryption to prevent malicious or accidental observation. Cryptography is

very complex and involved, but it for this paper it is enough to know that it works based on shared secrets such as keys or passwords. The key or password is used to transform the data into a scrambled and unreadable format for transmission. Once received, the shared secret or another prearranged method for reversing the original encryption is used to transform the scrambled data back into something meaningful.

Integrity is "…protection against unauthorized modification or destruction of information" (DOD 8500.1, 21). Integrity is implemented in computer networks by various means that involve calculations that cannot be reversed (known as "hashing") or by performing some calculation that can be compared by the receiver against the same calculation performed by the sender (known as a "checksum" or "bit checking"). These two methods allow the receiver to check his data against the sender's to ensure that nothing has been changed in transit.

Availability is "timely, reliable access to data and information services for authorized users" (DOD 8500.1, 17). Risk management ensures availability by using redundancy in design and operation and by using robust equipment. Managing networks can be a simple proposition for a small business or an incredibly complex one for larger, global organizations. Either way, if the network is unavailable any business, money is lost. Network operations are supported by back up sites, secondary systems, and other methods to avoid a failure preventing network access.

The final component of the model is the guarantee of authenticity. Authenticity is "establish[ing] the validity of a transmission, message, or originator, or a means of verifying an individual's authorization to receive specific categories of information' (DOD 8500.1, 17). Authenticity and authentication have a reactive role to stopping unauthorized personnel from gaining access to the network. Authenticity's proactive role is to ensure that information received is actually from the stated sender, and more simply, that the person on the other end of a connection on the network is who they say they are. Like Confidentiality, Authenticity is also protected with shared secrets and limited access. If a user is able to access certain resources, such as network assets, they are assumed to be authorized. In the same way, if they know the shared secret, such as a password or pin, they are authorized access.

CIAA is the backbone of network security. Every measure taken to protect the network is designed to protect one or more of the components of the CIAA framework. CIAA allows for a more detailed approach to risk management and helps to focus the efforts of network administrators.

3. The Three States of Information and What must be Secured

The protection of information on a network is the purpose of the CIAA framework. However, information can exist in three different states: processing, storage, and transit (Mairs, 127). Each state requires the protection of the CIAA. The state of the information and the particular threat for that stage will drive the security measures enforced by the network administrators. The threats are quite different for each state and can drive the budgets and security plans of an organization. The three states of information will be revisited in the next chapter when they are applied to covert social networks.

Information that is processing is normally being used by the computer. Computers take information from memory as needed and return it to memory when the necessary calculations are completed. Processing information must be protected because, while it is being processed, it is in danger of being changed or damaged (Mairs, 127). As the computer is reading the information out of memory, processing, and then writing it back to memory, the data is in danger. If the computer were to lose power or suffer a surge, the data it was processing and the memory locations that were being accessed will certainly be lost or need repair.

Data that is secured in memory is subject to unauthorized access or damage due to neglect, accidental overwriting, or component failure. Recent news reports about potential identity theft from stolen laptops and from accidental exposure on websites are timely examples of how stored data is often the most vulnerable state for information. Stored data's greatest vulnerability is the fact that humans easily forget the vast amounts of information that they have processed. Computers are designed to remember well past their own useful life. Systems with sensitive data may be exposed to threats because users forget what is stored on the systems. Often, stored data is a prime target due to the great return for hackers. For a simple analogy, it is a great deal easier, more efficient,

and more profitable to take money from a vault than it is to rob every person using a given ATM. The principle is the same for valuable data that is resting in storage.

Network security frequently focuses on information that is in transit. Information in transit is protected using encryption or by using secure channels that have very limited access. Network administrators are careful to ensure proper CIAA protections because data in transit is the most informative in terms of detecting a network attack. A hacker must move data on the network in order to achieve his objective. A virus must modify data for it to wreak havoc. In both cases, the transmission and reception of data can tip a network administrator that something is amiss. Additionally, the fact is that to hurt a network you have to gain access to it. Firewalls and intrusion detection systems block the majority of questionable incoming traffic before it can enter the network.

4. The Tactics of Network Security and How Networks are Secured

In a perfect world, computers would be designed from the ground up with CIAA in mind and there would be no threat from malicious users. However, as software and the protocols that are used to communicate were designed with a naïve belief that abuse was not a factor, security analysts and network administrators have had to adapt their thinking. Instead of secure computers, attention is now paid to creating a secure environment for the relatively weak machines. This adaptation has created the concept of the network perimeter.

The network perimeter is normally a logical division, but can also be physical or geographical. The concept is akin to building a wall around the network to keep the network safe and the hackers out. The network perimeter is where the bits in transit from the previous section are scrutinized for their activity. This is normally done through the use of firewalls and monitoring software. Firewalls that sit on the network perimeter examine inbound and outbound traffic for pre-defined patterns or characteristics. Once the traffic matches the programmed pattern, the firewall will block the traffic and can be set to sound an alarm. Recent adaptations such as Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) have expanded the concept of a firewall to add greater functionality.

In addition to a secure perimeter, network security implements a defense-in-depth (DID) approach with multiple layers of mutually supportive countermeasures. The perimeter does well in controlling the majority of hacking attempts, viruses, and other threats. However, if a threat does manage to penetrate the perimeter, there must be a way to detect the threat, neutralize it, and repair the damage. This is where personal firewalls, anti-virus, and anti-spyware factor into the equation. Each machine can be configured to act for its own protection. If a hacker gets into the network, he will still have to bypass additional security measures to do any damage to the machines. Furthermore, as each computer can be configured to allow only that traffic which is necessary for it to perform its assigned functions, the firewalls are normally more restrictive and quicker to alert if something is out of the normal. The anti-virus and anti-spyware programs are the last line of defense and scan the stored data on network computers to determine if an infection has occurred. Newer programs will scan all incoming data to ensure that they are not infected.

A final measure is logging network traffic and storing it for potential future use. If a network breech is discovered, logs can be examined to determine when the breech occurred and how it happened. This allows for a much more rapid repair of the network. Without logs, network administrators would have no idea how the breech occurred and would be forced to take the entire network down and start from scratch to be absolutely certain that all traces were removed from the network. Arguably, the greatest reason to log network traffic is to prosecute malicious users. Depending on how they are setup, logs can pinpoint the necessary information to be admissible in court and lead to convictions. There is also the added benefit of the psychological effect on users that know their activities are being monitored.

These combined tactics and control measures reduce risk as much as possible while still allowing the network to operate. However, the drawbacks are that they are expensive and they can demand a lot of processing power from the network. A compromise must be drawn between the protections necessary to keep the data safe and the value of the data in question. If the network's data is only worth $10,000, then it is absurd to pay for a firewall and antivirus suite that costs more than $10,000. There is hope that new operating systems like Microsoft's Vista, that are designed from the

beginning with security in mind, will put an end to the need for a network perimeter and DID, but it is highly unlikely that it will ever happen. Hackers are smart and they are motivated. New systems may stop old attacks, but the cat and mouse game will continue and hackers will find a new way to compromise these systems and networks.

D. HACKING AND THE COMPUTER NETWORK ATTACK (CNA) METHODOLOGY

The press has generally characterized hackers as the pimple faced teen in a dark room attempting to break into a bank's mainframe to steal enough money for a new Play Station 3. However, in reality many have simply taken the time to learn and explore computers to the point that they understand how the systems work (Verton, 2002). Although there are annual conventions where hackers meet to share their thoughts, there is a lack of organization in the community.

As network security emerged as a standardized field, a model was created that would explain the way that hackers penetrate a network. Although the origins are unclear, a model has been adopted that maps the process of hacking from its beginnings to exploitation. Here the model will be referred to as the Computer Network Attack (CNA) Model. A series of books designed to educate network administrators entitled *Hacking Exposed* has used the model since its first publication in 1999. Also, the International Council of E-Commerce Consultants has incorporated the CNA model into their Certified Ethical Hacker certification. The CNA Model has the following stages for hackers:

1. Footprinting
2. Scanning
3. Enumeration
4. Exploitation

Each successive stage provides a greater amount of information to the hacker and as the hacker knows more, he is able to do more damage and gain further access to network resources. The model is also applied by legitimate businesses that specialize in network security and testing, but instead of hacking it is termed "penetration testing" or "pentest." Either way, the model has provided the structure required to study CNA in a systematic way. Each stage of the CNA model will be examined in detail.

1. Stage 1 – Footprinting

Footprinting is the gathering of information on the target network. The information gathered ranges from the address of the target to the range of Internet Protocol (IP) addresses that are assigned to the target network by the Internet Assigned Numbers Authority. This information is the most general, and footprinting is normally passive. The target will not be probed in any way. The hacker uses publicly available information to gather as much detail on the network as possible.

Footprinting is a painful and slow process that requires a great deal of patience and perseverance. However, it is the first step because every bit of effort that is applied to footprinting will pay dividends in the later stages. To explain by analogy, McClure, Scambray, and Kurtz in their fifth edition of *Hacking Exposed* compare hacking to robbing a bank,

> For example, when thieves decide to rob a bank, they don't just walk in and start demanding money (not the smart ones, anyway). Instead, they take great pains in gathering information about the bank-- the armored car routes and delivery times, the video cameras, the number tellers, escape exits, and anything else that will help in a successful misadventure. The same requirement applies to successful attackers. They must harvest a wealth of information to execute a focused and surgical attack (one that won't be readily caught). As a result, attackers will gather as much information as possible about all aspects of an organization's security posture. Hackers end up with a unique *footprint*, or profile of their target's Internet, remote access, and intranet/extranet presence (6).

Footprinting uses public information as much as possible. Public databases are queried for network addresses, names, phone numbers, domain names (such as Google.com or Microsoft.com), and anything else that may be of use in later stages. It is important to note that footprinting is still a passive effort. Later stages will increase the visibility and the risk to the hacker, but during the first stage the hacker has done nothing to expose himself to possible detection.

2. Stage 2 – Scanning

Scanning is the first offensive stage and represents the first risk of detection to the hacker. Where footprinting passively gathered information about the network using publicly available knowledge, scanning actively determines the network's general structure, potential targets, and running programs that may be vulnerable to exploitation.

Again McClure writes that “If footprinting is the equivalent of casing the place for information, then scanning is equivalent to knocking on the walls to find all the doors and windows” (McClure, 42).

Automated tools will attempt to access different ports on the network. Using the *Hacking Exposed* analogy, the tools will try every port on a system as if it were a door. From the information gathered on the ports, educated guesses can be drawn as to what services are running and what systems are in use. This is where the openness of computer networking rears its head. Each operating system’s patterns are known and those patterns can tell the hacker what kind of machine is on the other side of the network connection. Each port that is open for traffic tells the hacker what kind of attack may be beneficial in the next phase. For example, if a network scan returns that port 80 is open, the scanned machine is most likely a web server and would be vulnerable to web server attacks. It is not a fool-proof method, but it is quite accurate and produces reliable results that require a minimum of tweaking.

Network scanning also gives the hacker a map of the network. Systems that respond to the scans tell the hacker that they are up and running. The hacker is able to map the network and the connections that exist between the machines. Again using well known protocol information and known best practices for network setup, the hacker can draw further conclusions about the functions of the machines. Some machines are more valuable targets than others. Computers that are storing passwords or acting as a router are good targets. Best practices and convention dictate that routers are always given the first IP address in an IP space. In other words, when a hacker gets a reply to his scan from a machine whose last digit is a one (e.g, 192.168.1.1) he can reasonably assume that the machine is a router. Routers act as connectors in the network and are important to mapping as well as further exploitation.

3. Stage 3 – Enumeration

Scanning tells the hacker what machines are listening on the network and what potential services they are running. Enumeration verifies that the hacker’s theories about services are correct. Now the hacker seeks to know exactly what versions,

configurations, and security measures he is facing "…[hackers] typically turn next to probing the identified services more fully for known weaknesses, a process we call *enumeration*" (McClure, 78).

Hackers will use techniques to determine the exact versions of software that they are facing. Each software version is slightly different from the previous version. New versions are released due to improvements in the software, but they are also released due to known security holes that need to be patched. If a hacker can tell which version of a software package is in use, he can search for security holes to exploit.

In enumeration, networks begin to lose private information. Scanning is an active attack on the network, but it does not directly aim to gather information on people or the data stored on the network. Scanning learns about the computers. Enumeration aims to learn more about the running systems and the human-machine interface. User names, the probable location of password files and the exact location of sensitive data are probable returns for the hacker during enumeration.

Enumeration is most effective when network administrators do not properly configure their systems. Often when connecting to different machines or services, extraneous information is given away. Network security battles this with the Principle of Least Privilege or POLP, and it basically states that users only need as much information as necessary to complete their tasking, no more, no less. However, in a network with hundreds or thousands of machines, it is easy to forget to sanitize the extra information.

4. Stage 4 – Exploitation

Exploitation is the final stage of an attack. This is when all of the previous stages come together allowing the hacker access to the network. Obviously, this is the worst possible situation for a network. Any data that was protected has now been exposed. Even more damaging than what was actually done to the network is the potential damage. Administrators may not be able to determine the true extent of the hack and will be forced to purge any system that the hacker *may* have had access to.

Exploitation relies on the hacker's creativity to gain access to the computer network. Although the network's operations and configuration are known, the penetration method and target selection are a matter of experience, skill, and desired

outcome. Just as with destabilizing a covert network, exploitation in CNA is the point where the systematic approach must pause for the inclusion of creativity. However, the hacker basically has three avenues to pursue.

The most traditional exploitation is the system exploit. This is when the hacker specifically targets a certain machine based upon its value or upon its poor security. If the system is one that will guarantee great rewards, a hacker may focus on that machine until a hole is found. Another approach is to look for the "low hanging fruit" and attack those machines that have weak security or improper configurations. Once the weak machine is penetrated, it can open the door to the harder machines. Whatever the case, system exploitation works because the hacker determines what software is running and then uses a known security flaw or a new flaw that he discovered. This is the reason that network security focuses on software "patches" or code that fixes insecurity in the original software. Microsoft uses what is known as "Patch Tuesday" to release new patches for their operating systems and other software.

Another approach to exploitation is to attack the network traffic. A hacker will watch network traffic to determine who is talking and what they are talking about. Once the network session is identified, the hacker will carefully pick his time, jump in the middle of the session performing a "session hijack." Session hijacking is dangerous because the hacker appears to be an authorized host, and there is no need for the hacker to authenticate himself (Whitaker & Newman, 127). Once the session is hijacked, the hacker impersonates both ends of the session to the other. This is known as a "man-in-the-middle" attack and allows the hacker to see every piece of data transferred between the two machines.

The last form of exploitation is attacking the user. Attacking the user allows a hacker to skip the three previous stages and jump right to exploitation. There are multiple ways to attack a system user. The most common is to trick the user into opening a file or an email attachment that contains a malicious program. These programs are called "Trojan horses" or "worms." There are subtle differences between a virus, a worm, and a Trojan horse, but they all aim to attack the user to gain access. The more skillful approach is to attack the user with a technique called "social engineering." Social

engineering is “the act of tricking another person into providing confidential information by posing as an individual who is authorized to receive the information” (Harris, 967). Using social engineering, a hacker will call a user and pretend to be from the Information Technology (IT) department at the business. He will attempt to say the right things to fool the user into giving up their password or other sensitive data. Social engineering is a skill that takes practice and a little bit of knowledge, but it can pay great dividends.

E. CONCLUSION

CNA, network security, and computer networks have lead an iterative process of growth and change. The nature of computer networks drove the need for security to stop hackers. Hackers find holes in the systems and necessitate changing security tactics and modifications to the computer networks function. The three components are constantly changing and adapting to keep the balance of security and functionality. The CNA Model is the result of this give-and-take.

With a clear understanding of the CNA model, it is now possible to compare computer networks and social networks. The validity of using CNA as a framework for the SNA of covert networks will be examined to determine if the CNA model can improve the SNA process.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. HACKING A SOCIAL NETWORK

A. INTRODUCTION

Up to this point, social networks and Social Network Analysis (SNA) have been introduced and examined. The shortfalls of SNA when applied against covert social networks were made apparent. Computer networks, computer network security, and the Computer Network Attack (CNA) model were covered in chapter two. The necessary background information has been covered to address the purpose of this paper: to determine if the CNA model can improve SNA when applied to a covert social network, and how applicable the model is to destabilizing a covert social network.

This chapter will examine how computer networks and social networks compare to one another. They are alike in just as many ways as they differ, but the differences are great enough that the CNA model can only be used as a general, strategic guide for destabilization instead of a tactical aid to exploitation. Finally, the CNA model will be examined to demonstrate its potential for improving the defeat of covert networks, and the model will be applied to SNA to demonstrate how it can be used to integrate SNA into an offensive attack against covert networks.

B. COVERT SOCIAL NETWORKS AND COMPUTER NETWORKS COMPARED

1. Similarities

With either covert social networks or computer networks the focus of interest is on the network as a whole instead of the people or machines that compose the network. We recall that SNA was developed to study the connections between people instead of studying the people who have connections. Computer networks are studied in the same way. The focus is not on the machines themselves, but on the network in its aggregate. The focus on connections is a basic premise shared by the separate fields. These fundamental frames-of-reference also created the same focus on network metrics. For social networks, metrics include such things as centrality, transitivity, and prestige. For computer networks, metrics examine network data such as throughput, latency, and loss.

For either social networks or computer networks, the metrics used to understand the network are not based on the nodes of the network, but are measurements of how the network interacts with the node.

However, probably the most obvious similarity between computer and covert networks is the fact that both share a network structure. Both computer and social networks are characterized by the interconnection of nodes that share a common relationship of some kind. The nodes in question are of relatively equal rank. For computer networks, nodes are connected by the need to share information or by the need to use shared resources like printing or Internet access. Covert social network relationships are also connected to share information. All interconnection is to share information at some level, regardless of the relationship, and a terrorist's ties to his family, friends, or other terrorists exist out of the need or desire to share information. As for the need for a shared resource such as printing or Internet access in a computer network, a simple example for a terrorist network is access to a bomb maker. The bomb maker is a resource that the other members of the network require access to and it creates a need for interconnection.

To examine the shared structure even further, covert networks can be examined using the topologies of computer networks. The most common structure that is encountered by counter-terrorist forces is the cellular structure that is used by experienced groups to avoid detection and/or minimize damage if members are captured. For a computer network, the cellular structure is identical to a tree topology with star subnetworks. The main line is connected to the central machine and the central machines have other computers connected to it. This is the same as the independent cell that has a single leader with anonymous access in some way to the group's leadership and planning. The terrorist cells are small and compact with limited connections between members. However, imagine the opposite structure where all network members are interconnected to other members. This is the same as the mesh topology that is now becoming more commonplace in wireless networks. Certainly, terrorists do not use a mesh topology due to the fact that the capture of even one network node would present a grave security problem, but the illustration is effective in proving that the methods used to describe the structure of computer networks can be applied to covert social networks.

Another way to compare the two networks is by using the general divisions of Local Area Network (LAN) and Wide Area Network (WAN) used in computer networks. The small cells of terrorists connected with the fewest possible number of connections forming a dyad, triad, or a subgroup is an ideal parallel to a LAN. Additionally, with a LAN there is normally a central management that coordinates and directs, and this is the cell leader in the case of a covert network. Pulling back to examine the cell in the larger organization of the entire terrorist social network creates the interconnection of multiple LANs and is the textbook definition of a WAN in computer networks. Just as the distinction between a LAN and WAN is important to understanding computer network behavior, the division is also important if applying the concept to covert networks. WANs are characterized by a loose management and less restrictive policies that tend to serve as guidelines instead of directives. The covert network example is that Al Qaeda may issue a statement to their operatives that they should attempt to incite factional violence, but their statement will only serve as a suggestion to the cell leader in Iraq as enforcement is impossible due to the distance between the Al Qaeda leadership and Iraq. The LAN, or cell, will enact their own polices and procedures that are much more restrictive due to the proximity to the cell members.

Covert networks cannot eliminate risk and must practice risk management using the same rationale as computer network security. Covert networks must operate with accepted risk. Terrorists cannot eliminate the forces that are working against them. To remove all risk would require the killing or neutralization of all government forces or a complete cessation of all illegal activity. Although the elimination of all government forces would be ideal for a terrorist, it is obviously impossible. Likewise, ceasing attacks and laying down arms is not an option if the group is to attain its objectives through violence. The result is a need to operate with some assumed risk and minimize the likelihood of exposure, capture, and death to the point that the covert network can operate and still achieve their goals. The security models for computer networks differ from covert networks, but the shared characteristic is that both networks function in a threat environment that requires security to be an integral part of operations.

Communications in a covert network can be synchronous or asynchronous. When communicating in a covert network, options include face to face meetings, dead drops,

couriers, runners, phone calls, email, web pages, etc. Some of these methods are synchronous (phones and meetings) and others are asynchronous (dead drops, email, runners, web pages). When talking in person or on the phone, the timing required to carry on a two-way, simultaneous conversation is done automatically as long as one person doesn't interrupt or talk over the other person. When communicating using asynchronous methods, there is no need to coordinate as one person sends their message and waits for the return like a tennis match.

For a terrorist, asynchronous messages are safer than synchronous because they reduce the risk of detection, but they also come with the drawback that a reply message may come almost instantly or never at all; this makes planning and coordination difficult. A danger is both parties in the communication creating their own understanding of what was being said. One message may say, "Attack at noon," and the next message may say, "Change attack to 1230." An asynchronous communication would leave both sides in confusion until some type of acknowledgement was received that the attack time was changed and that all players are working from the same plan. This demonstrates why asynchronous communications are used for simple orders or issues that are not time critical. Synchronous communications are used when planning an operation, clarifying important details, or in the final phases of an operation to ensure that all actors are properly coordinated. Knowing the two forms of communication can be useful for determining the likelihood of an attack, but it comes with the caveat that it is unpredictable. Understanding the difference in synchronous and asynchronous communications can help the network analyst determine that a change in the covert network is occurring, but the reason for the change and what it means to friendly forces is something that would be determined by other means.

Information in a covert network can be in one of the three states outlined in the computer networks chapter: processing, storage, or transit. Information that is processing is being used by the group to achieve its mission. It is information that is important to the group in the immediate present. For example, if a cell is waiting in ambush for a target, the processing information is the tasks and roles that each member of the cell is to perform when the ambush occurs. When the signal is given, the processing information is acted upon and more information begins processing, such as their escape

plan. Stored information is that information that is known to the terrorist, but is not being used. The location of the safe house, the members of the cell, and their normal communications methods are examples. Information that is in transit is communication that carries current and time sensitive instructions or details. Using the previous ambush example, the information in transit is the order to execute the ambush.

For friendly forces fighting a covert network, information in transit and information in storage are the two most important. This is the same for computer networks. Information in transit is paid the most attention in computer networks because it can indicate an attack is ongoing or is imminent. The same is true for a covert network. The frequent reports of "increased activity" or "increased chatter" that the DOD uses to raise the terror threat warnings are proof that information in transit is an important indicator of covert network activity. Stored information is important to tracking, mapping, and neutralizing a covert network. Capturing a cell member and using his stored information allows friendly forces to determine the necessary details to continue their efforts at stopping the network.

The final similarity between the two networks is one of strategy. Computer networks and covert social networks function on the Principle of Least Privilege (POLP) mentioned in the previous chapter. POLP is frequently encountered in another expression, "need to know." Covert networks survive based on maintaining a "need to know" concept of operations. There is a strict compartmentalization of the network's members and only those that need to know information are given access. POLP is used by computer network administrators to keep hackers from gaining information that is potentially damaging.

2. Differences

First and foremost, the openness and standardization of computer networks is absent in covert networks. Covert networks have a standard way that they communicate, but it is not known openly. What is more, the standard means of communication can be quickly changed in the event of a security breech. Covert networks intertwine obscurity into every aspect of their operations. The two networks have opposite requirements for longevity. To remain a functioning network, a computer network must maintain

predictable operations and a predictable configuration. If a covert network becomes predictable, it increases its risk. Therefore a covert network seeks to maintain secrecy at all times for its own protection.

The strategic approach of the Confidentiality, Integrity, Authentication, and Availability (CIAA) Framework used to protect computer networks is replaced by trust in covert networks. Computer networks itemize the characteristics of information that must be protected to ensure a secure operating environment. Covert networks do not itemize, but instead take a general approach that is all encompassing. Confidentiality is protected through trusted couriers carrying messages, or basic code words and other simple methods, which are only known to trusted contacts to prevent observation. Authenticity is protected by only using the trusted members of the network or other contacts that are vetted by those that are already trusted. Integrity must rely on trusted couriers or direct contact to prevent the alternation of communication information. Finally, the availability of network access is not protected by trust, but is protected using tradecraft such as frequent relocation, redundant communication channels, and other measures (Sageman, 2004 and Arquilla, 2001). Covert networks benefit from their simplicity when compared to even the smallest LAN and are able to operate securely using a less scientific or dogmatic approach.

Covert networks use compartmentalization to protect an ambiguous network perimeter. Computer networks may also suffer from borders that are confused and shifting, but compared to the challenges outlined in SNA in creating a perimeter for social networks, they are much more obvious for computers. The use of firewalls and a defense-in-depth approach of computer networks centralize resources behind a curtain of protection. This is infeasible for covert networks and they are forced to rely on alternate means. They distribute manpower, resources, and knowledge as much as possible to minimize damage in the event of a compromise. Covert networks cannot build a wall around their network to protect it like a computer network does as a part of standard practice. This is one of the reasons that covert networks are hard to fix and finish. The sacrifice for the covert network is that communications suffer from speed and reliability problems.

There are no governing or regulating bodies that dictate the inner-workings of a covert network. There is nothing similar to the Internet Assigned Numbers Authority (IANA) in computer networks that instructs a covert network how to carry on communications or what protocols should be used to share information. The closest that a covert network may get would be a training division that attempts to standardize the network's operations. The lack of standardization creates a learning curve for new network members and for law enforcement attempting to penetrate a network. Although there are general trends that are common to covert networks and they often respond in typical ways when pressured by law enforcement, it cannot be assumed that they are enough alike to not require independent study. Each cell or group may function differently and necessitates a lengthy process of study to understand the intricacies of the group. The positive of this lack of standardization is that it also forces the covert network to rely on simple methods that change infrequently. Attempting to modify the network's method of operations or communications is not an easy task and must be done slowly and carefully to transfer all of the network's members to the new system. With computer networks, the IANA can release a new regulation and every computer wishing to connect to the Internet knows what is expected and how to meet the regulation. The lack of regulation is a double edged sword.

Finally, along with a lack of regulation and standardization comes the lack of a unified terrorist topology or standard operating procedure. Where there are only a few options for operating systems in computer networks that can focus an attempted attack and make the selection of exploitation method a simpler affair, there are no parallels in a covert network. Each terrorist, cell, and group is different. They follow common trends and generally organize along the same lines, but there are no guarantees. This creates a challenge for the analyst attempting to destabilize a covert network. The creativity needed by the hacker when exploiting a computer network was at least bounded by the limited possibilities in computer network configuration. This is again a double edge sword as the bounds that limit a hacker's ability to find a way to exploit the network do not exist for an analyst attacking a covert network.

3. How They Compare

The similarities and differences are marked between computer and covert social networks. The direct similarities in security, risk, structure, information sharing, and outsider threats allow the application of the CNA model to a covert social network without drastic modification. However, it is clear that the lack of predictability, standardization, and regulation force the CNA model to remain at the strategic level functioning as a guide to exploiting a covert social network. The CNA model applies to the general process of destabilization. Attempting to take the model down to a more granular approach causes it to fail. For example, the CNA model for computer networks uses automated programs to scan computers on the network for open ports. Covert networks do not have open ports that are assigned by standardizing bodies, and therefore, there are no automated programs to scan a covert network. The CNA model can be used in destabilizing a covert social network, but only when it is used to organize SNA into a systematic, phased approach. The CNA model cannot be used as a direct comparison of hacking a computer network to destabilizing a covert network because of the differences between the networks.

C. APPLICATION OF THE CNA MODEL

1. The Validity of the CNA Model

When applying the CNA model to SNA, the primary purpose is to supplement the SNA process using a systematic approach in order to reduce errors and increase confidence in destabilization strategies. SNA alone is useful and can possibly help the forces battling a covert network gain a deeper understanding of the network and may provide insights as to the best way to bring the network down. However, as was discussed at length in the SNA chapter, the absence of a standardized approach to conducting SNA creates opportunities for mistakes and faulty assumptions.

It is clear that the CNA model can be applied as a general aid to destabilizing covert social networks, but the question remains if it will really help reduce the inherent error in analyzing a covert social network. It is feasible to take a given social network and conduct two different network analyses, one without applying the CNA model and one with the model, and then compare the results. However, the volatility of social networks prevents such a comparison yielding any reliable results.

The difficulties SNA faces when applied to a covert network may be minimized when the process is standardized. Although untested, it is reasonable to assume that the standardization of SNA will improve accuracy due to the possibility of adding peer review and by limiting assumptions. Each phase can be limited by a confidence level that prevents further planning until that level is achieved. For example, an operation to capture a wanted terrorist can be held until the analyst employing the CNA model can produce a 90% confidence level that the person in question is the right target for the right reasons. Although subjective, it does provide a chance to review the process at each phase.

Additionally, at the conclusion of each phase, the assumptions made during the analysis can be itemized for peer review. Any underlying assumptions can be challenged, and by bringing them to light, the analysis can be further improved. SNA tells the analyst nothing about the network. SNA only presents a matrix or a graph that attempts to model the network's relationships. Every value or characteristic assigned to the social network is based on the interpretation of the analyst. Interpretation of something as complex as a social network creates a high probability for faulty assumptions. It is necessary to challenge these assumptions to improve SNA and to take SNA towards an offensive tool instead of a descriptive one. To provide an example of challenge assumptions, targeting a node because it is the most central to the network assumes that the node is critical to the operation of the network and it also carries an underlying assumption that centrality is the network characteristic that deserves to be targeted. Peer review by other analysts or decision makers can examine the first assumption that the node is critical to network operation and the second assumption that centrality is the right criteria to choose for destabilization. This peer review should increase the accuracy and confidence in destabilization efforts.

The chapter on SNA concluded that the challenges of applying SNA to covert networks cannot be completely overcome due to inherent data errors. If the data cannot be improved, then to increase confidence levels in network analysis, the process that is used to examine the data must be improved. The CNA model provides an excellent starting point for improving SNA.

2. Application

Computer networks and social networks are similar enough to apply the CNA model, and it is reasonable to believe that the CNA model is qualitatively valid, so the final question that remains is how the model can be applied to destabilize a covert social network. Before the model is applied, a change in mindset may be required by a network analyst. SNA methods and techniques were created to support the academic study of open networks. However, to attack and destabilize a social network using the CNA model, the analyst must move from the academic mindset to an offensive mindset that is aimed entirely at exploiting the social network. SNA is a descriptive tool, and ideally, the use of the CNA model can improve the offensive capabilities of SNA when applied to covert networks. Not only can CNA provide structure for the analyst and increase the confidence levels of destabilization strategies, it can also help to guide the analyst towards the aggressive and offensive mindset necessary to topple a hardened covert network.

The analyst must become a social network hacker. The analyst must also recognize that the goal at the beginning of a destabilization effort should not be to topple the entire network in a single action. The goal is to find a foothold for exploitation. Once the foothold is secured, the CNA model can be applied repeatedly to hopefully lead to the eventual destabilization of the covert network. Theoretically, little by little, the network can be toppled, but attempting to do it from the outset focuses on the forest instead of the trees. Actually bringing the network down will depend on a great number of factors, but beginning a destabilization using smaller maneuvers with a cumulative effect should increase the predictability of the results and provide a controlled approach.

The destabilization results of applying the CNA model to a covert social network will surely vary depending on a multitude of factors such as the skill of the analyst, available information, and the objectives of the analysis effort. What can be stated with certainty is that applying the CNA model to a covert social network will not be as successful as applying it to a computer network. The differences in complexity and predictability between computer networks and social networks prevent correlated results. The CNA model points directly to the holes in a computer network, but when used to attack a social network, it merely provides structure and standardization to SNA. When

the CNA model is applied to a computer network, any vulnerability in the network will emerge as a hacker progresses through each of the four stages. This is not necessarily true when the model is applied to a social network. The analyst will still have to determine where to attack the network instead relying on a program that finds the network's weaknesses. Social networks are more complex than a simple port being open or an improperly secured computer service. The CNA model is designed to filter out vulnerabilities in comparably simple computer networks. Unlike computer networks, there is no guarantee that the CNA model will find any of the vulnerabilities in a covert social network. This is a limitation of the CNA model when it is applied to covert networks and it is a point that must be remembered by the analyst.

One of the benefits of applying the CNA model is that its phased approach marries well with the phased planning done by the military. The staged process that the CNA model uses is ideal for incorporation into the Joint Planning and Execution System (JOPES) or Military-Decision Making Process (MDMP). Both of these planning processes use a multi-phase approach to move from initial planning to the final stages and clean up of a military operation. Each stage of the CNA model can match the phases of military planning.

To further demonstrate the application of the CNA model, each phase will be examined in detail with possible arrangements of how the model can assist in the application of SNA against a covert network.

a. Footprinting

When applied against a covert social network, footprinting is the collection of open source information and closed source intelligence about the network and its likely operations. The analyst uses news reports, public records, the Internet, intelligence reports, and any other sources of information to gain a tentative picture of the network. In the first stage, the data can be of questionable veracity or from unreliable sources. The idea is to gather everything and anything possible to prepare for the following stages where it will be examined for quality.

After the data is collected, the analyst sorts through the information and attempts to determine a basic understanding about the network. This is the "rough draft"

of the SNA process that will be updated and changed. The likely topologies for the network are examined and key players are identified. Footprinting requires the analyst to make assumptions about these network members and the organization of the network as a whole. Each assumption can be analyzed to ensure that it is reasonable and also help identify missing information that is required to verify or dismiss the assumption. Assessments on network metrics such as centrality are not measured, but are at a minimum, starting to show themselves to the analyst. Initial projections and destabilization strategies can be presented, but attempting to plan anything this early in the process would be a waste of resources as the initial understanding of the network may reflect very little about the actual organization.

b. Scanning

Scanning a covert network, just like a computer network, is the first active phase in the attack on the network. The objective of the scanning phase is to start eliminating some of the assumptions from the footprinting stage, improve the projected network model for topology and personnel, and begin active collection of new intelligence. Scanning is the first active phase because new intelligence must be collected to determine how the dated facts used to in footprinting actually reflects the current network. The dynamic nature of covert networks requires continual intelligence to maintain network models with as little error as possible. Scanning is the beginning of the intelligence collection. Initial projections for network metrics can be generated as well as tentative destabilization strategies. The metrics and strategies will have low confidence levels until the data is further analyzed, but scanning represents the first time that the analyst can reasonably determine what the network looks like and how it operates. This is the information that is required to give commanders options to topple a network.

c. Enumeration

Enumeration incorporates the cornerstones of SNA. The data collected in footprinting and updated in scanning is used to apply SNA metrics and analysis tools. The data is combined into sociomatrices and sociograms. The frame of reference for the analysis is finalized and determines the number of modes, the coding schemes, and the network boundary. Network metrics are measured and their validity is the highest it will

ever be. Any previous assumptions are revisited and dismissed, or are noted for inclusion in the final confidence level calculation.

Enumeration is concerned with pattern detection. The first of the two pillars of SNA emerges when the data is fully examined for the first time. The previous two stages projected the topology and network roles and relationships, but enumeration specifically addresses these network characteristics to provide the definitive projection with as little error as possible. Enumeration focuses on who is connected to whom. Again, this may require assumptions on the part of the analyst, but each assumption can be listed and can reference the supporting evidence for each one. These assumptions can be examined by analysts or superiors not involved in the network analysis to determine if they are logical and reasonable.

d. Exploitation

This is the stage that requires the most skill and creativity from the analyst. The strategies outlined in the SNA chapter of nodal removal and nodal manipulation must be used against the correct node. The choice of destabilization tactic is situation dependant and incorporates the uncertainty of the previous stages in the CNA model. Accuracy is vital and is the reason that each preceding stage must contain as few errors or unqualified assumptions as possible as their effects are cumulative.

Exploitation begins with pattern interpretation and ends with the offensive measure used to destabilize the network. Exploiting a covert network requires the analyst to interpret the patterns detected during enumeration to find the weakness that will provide the foothold for further exploitation. Once the causes of the organization of relationships are determined in the network, the analyst is able to predict the first-order effects on the network with a reasonable level of certainty.

Targeting is an important component of exploitation. There are two options. The first is to let the network analysis determine the targeted node. This has the advantage of allowing the analyst to use creativity to create the optimal destabilization plan by choosing the nodes that will cause the greatest damage, or choosing nodes that will lead to follow-on attacks that will create the greatest damage. The second option for targeting in exploitation is to already have the node targeted and then determine the

network's response. Instead of focusing on the network and how to destabilize it, the focus in on the node and what the loss will mean to network operations. The model can be used to determine network effects, assuming that footprinting, scanning, and enumeration are all complete. An example would be an order to capture a High-Value Target (HVT) from a higher authority. The analyst would be able to provide reasonable predictions, at least first order effects, on how the network would respond to the removal of the HVT. Another example would be a chance killing of a network node and determining how the network will respond.

The exploitation stage is the final step in a long chain of careful analysis, and the model cannot be used out of order. Users of SNA may wish to rush into exploitation without an appreciation of the buildup that is necessary. There must be an understanding of how the network functions and why it functions as it does. Without this understanding, selecting a target for removal or manipulation will produce unpredictable results and may do more damage than good. The CNA model provides a form of a cost-benefit analysis. Commanders must at least know the price that they potentially pay for independent targeting. The analyst is able to provide reasonable conclusions about what the network may do in response to an offensive action. The bottom line is that when the CNA model applied to SNA, each stage allows for a chance to pause and examine the assumptions made or assumptions dismissed—ensuring that the process produced the highest confidence level possible.

V. CONCLUSION

A. REVISITING THE OBJECTIVE

The objective of this paper was to determine if the Computer Network Attack (CNA) model can improve the use Social Network Analysis (SNA) against a covert social network. The conclusion is that the CNA model is applicable to covert network and it is reasonable to believe that it will reduce error in conducting SNA. Computer networks and social networks share enough similarities that the CNA model can be applied as a strategic guide to destabilization without major modifications to the model. Ideally, the integrated process outlined in the final chapter which intertwines the current approach to SNA with the CNA model would be used by forces battling a covert network and desiring to increase the confidence levels in their predictions. With the recent incorporation of SNA into military doctrine, SNA it its present form will certainly have to be modified to meet the needs of the military user engaged in offensive operations. Unfortunately, it will never be the predictive tool that can definitively lead to the destruction of a covert network. That being said, the CNA model is the perfect beginning of the expanded use of SNA as an analysis tool, and it has a future in standardizing and improving the process of exploiting a covert network.

B. SUGGESTIONS FOR FUTURE RESEARCH

The following areas are starting points for future research based on this paper:

- Apply the CNA model to an actual covert network and specifically propose how the model can destabilize the network.
- There exists the possibility for a final stage to be added to the CNA model after exploitation to handle feedback and model correction.
- Write an annex to the new counterinsurgency manual, FM 3-24, that details how the CNA model can be added to current doctrine.
- Parallel the CNA model and covert network exploitation with the staged military planning found in the Joint Planning and Execution System (JOPES) or Military-Decision Making Process (MDMP).
- Further refine the model with expanded input from computer scientists, network administrators, sociologists, and military or law enforcement personnel that have experience using SNA against covert networks.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Barabási, Albert-Laszló. *Linked : How Everything is Connected to Everything Else and what it Means for Business, Science, and Everyday Life*. New York: Plume, 2003.

Barabasi, Albert Laszlo, and Eric Bonabeau. "Scale-Free Networks." *Scientific American* 288.5 (2003): 60.

Brenton, Chris, and Cameron Hunt. *Mastering Network Security*. Second ed. Alameda, CA: Sybex, 2003.

Carley, Kathleen, Ju-Sung Lee, and David Krackhardt. "Destabilizing Networks." *Connections* (2001): 31-4.

Carley, Kathleen, Jeffrey Reminga and Steve Borgatti. "Destabilizing Dynamic Networks Under Conditions of Uncertainty." International Conference on Integration of Knowledge Intensive Multi-Agent Systems. 2003. Accessed July 16, 2006 <http://www.casos.cs.cmu.edu/publications/papers.php>.

Carley, Kathleen. "Destabilizing Covert Networks." *Computational & Mathematical Organization Theory* 12.1 (2006): 51-66. Accessed July 16, 2006 <http://springerlink.metapress.com/>.

Carley, Kathleen. "Destabilizing Terrorist Networks." *Proceedings of the 8th International Command and Control Research and Technology Symposium*, 2003. Accessed July 16, 2006. <http://www.casos.cs.cmu.edu/publications/papers.php>.

Carrington, Peter J., John Scott, and Stanley Wasserman. *Models and Methods in Social Network Analysis.* Cambridge; New York: Cambridge University Press, 2005.

Cross, Robert L., and Andrew Parker. *The Hidden Power of Social Networks : Understanding How Work Really Gets Done in Organizations*. Boston, Mass: Harvard Business School Press, 2004.

Degenne, Alain, and Michel Forse. *Introducing Social Networks*. London: Sage, 2004.

Department of the Army. *FM 3-24/MCWP 3-33.5: Counterinsurgency*. Washington, 2006.

Department of Defense. *DOD Directive 8500.1 Information Assurance (IA)*. 2002.

Farley, Jonathan. "Breaking Al Qaeda Cells: A Mathematical Analysis of Counterterrorism Operations (A Guide for Risk Assessment and Decision Making). *Studies in Conflict and Terrorism* 26 (2003): 399-411.

Fukuyama, Francis and Abram N. Shulsky. *The 'Virtual Corporation' and Army Organization*. RAND ARROYO CENTER SANTA MONICA CA, 1997.

Goleniewski, Lillian. *Telecommunications Essentials: The Complete Global Source for Communications Fundamentals, Data Networking and the Internet, and Next-Generation Networks.* Boston: Pearson Education, 2003.

Harris, Shon. *CISSP Exam Guide*. 3rd ed. New York: McGraw-Hill/Osborne, 2005.

Krebs, Valdis. "Mapping Networks of Terrorist Cells." *Connections* 24.3 (2002): 43-52.

Mairs, John. *VPNs: A Beginner's Guide*. Berkeley, CA: McGraw Hill/Osborne, 2002.

McClure, Stuart, Joel Scambray, and George Kurtz. *Hacking Exposed: Network Security Secrets and Solutions*. Emeryville, CA: McGraw Hill/Osborne, 2005.

R. Albert, H. Jeong, and A.-L. Barabási. "Error and Attack Tolerance in Complex Networks." *Nature* 406 (2000): 378-81.

Raab, Jorg and H. Brinton Milward. "Dark Networks as Problems." *Journal of Public Administration Reasearch and Theory.* 13.4 (2003): 413-439.

Rand Corp. Santa Monica, CA, et al. *Countering the New Terrorism.* 1999.

Rand Graduate School Santa Monica, CA, and Kiser D. Stephen. *Financing Terror: Analysis and Simulation to Affect Terrorist Organizations' Financial Infrastructures*, 2004.

Rand National Defense Reasearch Inst. Santa Monica, CA, John Arquilla, and David Ronfeldt. *Networks and Netwars. the Future of Terror, Crime, and Militancy.*, 2001.

Sageman, Marc. *Understanding Terror Networks.* Philadelphia: University of Pennsylvania Press, 2004.

Scott, John. *Social Network Analysis: A Handbook.* Second ed. London: Sage, 2005.

Verton, Dan. *The Hacker Diaries: Confessions of Teenage Hackers.* Berkeley: McGraw-Hill, 2002.

Wasserman, Stanley, and Katherine Faust. *Social Network Analysis: Methods and Applications.* New York: Cambridge University, 2005.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Head
   Information Operations and Space Integration Branch
   PLI/PP&O/HQMC
   Washington, D.C.

4. Professor D.C. Boger
   Naval Postgraduate School
   Monterey, California